Development and Evaluating Self-Health Literacy Assessment in Pre Service Science Teacher: Analysis Rasch Model and SEM-PLS

Annisa Nur Khasanah[1,2]*, Dadan Rosana[1], Insih Wilujeng[1], Arifin Septiyanto[3], Muhammad Nur Hudha[2]

[1]Universitas Negeri Yogyakarta, Colombo St. No. 1, Sleman, Special Region of Yogyakarta, 55281, Indonesia

[2]Universitas Sebelas Maret, Ir. Sutami St. No. 36A, Surakarta, Central Java, 57126, Indonesia

[3]Universitas Islam Negeri Sunan Ampel, Ahmad Yani St. No.117, Surabaya, East Java, 60237, Indonesia

*Corresponding author, email: annisa464fmipa.2022@student.uny.ac.id

doi: 10.17977/um065.v6.i2.2026.10

Article history
Submitted: 5 January 2026
Revised: 13 February 2026
Accepted: 14 February 2026
Published: 15 February 2026

Keywords
Pre service science teacher
Rasch model
Self-health literacy
SEM PLS

Abstract
The study was conducted to develop and evaluate a Self Health Literacy Assessment Instrument using the Rasch Model and Structural Equation Modeling with Partial Least Squares (SEM-PLS) in Pre Service Science Teacher. The instrument was designed to evaluate individuals' ability to access, understand, process, and apply self health information. Development proceeded through three stages: domain identification, item generation, and instrument construction. Pilot testing involved 211 pre service science teacher (25 male, 186 female). Twenty five items were initially administered; nine were removed due to low factor loadings and Average Variance Extracted values below recommended thresholds. Rasch analysis indicated robust psychometric performance, with person reliability of 0.89 and item reliability of 0.98. Infit and Outfit Mean Square statistics near 1.00 evidenced item fit, and point measure correlations indicated satisfactory discrimination. Complementary SEM-PLS results supported construct validity, with convergent validity acceptable (AVE = 0.550–0.583) and composite reliability adequate (CR = 0.805–0.865). A significant predictive path was identified between Understand and Process (β = 0.536). The finalized instrument contained 25 items by Rasch and 16 items by SEM-PLS across four dimensions. Overall, the measure is reliable and valid for assessing self health literacy and suitable for application in education and intervention research.

1. Introduction

Health is seen as a crucial element in the social aspects of sustainability at various levels whether individual, community groups, or organizations (Reyna-Castillo, Pulgarín-Rodríguez, et al., 2022; Reyna-Castillo, Santiago, et al., 2022). This is especially important in formulating self-health guidelines that aim to maintain health and well-being sustainably. Without conditions that support public health, long-term development is not possible. Therefore, efforts to design intervention guidelines for sustainable health and well-being are closely related to the concept of self-health (WHO, 2022). Health self-care is the habit and skill that helps maintain good health and a positive attitude to face life's challenges and continue to thrive (Wise et al., 2012). Self-health not just about maintaining physical health but also includes the attitude and desire to protect oneself (Kissil & Niño, 2017). This can be done by thinking about oneself, meeting personal needs, and seeking help or resources that support health and well-being (Colman et al., 2016; Pakenham, 2017). There are many ways to practice self-health, such as maintaining hygiene, eating healthy food, and doing activities that make life more comfortable (Urpí-Fernández et al., 2020). In addition, self-health also helps a person feel happier and more prosperous, for example by managing emotions, calming the mind, increasing self-awareness, doing spiritual activities, and interacting with nature (Butler et al., 2019; Corral-Verdugo & Frías-Armenta, 2016). In this context, self-health literacy assessment plays a crucial role in enabling individuals to identify potential health risks, monitor their health condition regularly, and take necessary preventive or curative measures.

Self-health assessment not only serves as an early detection mechanism but also as the foundation for holistic self-care practices, ultimately contributing to increased happiness and overall well-being (Acoba, 2024; Martín-Rodríguez et al., 2024). By regularly monitoring health conditions and having awareness of one's own state, one can more easily control emotions, calm the mind, and carry out activities that support mental and spiritual health. This contributes to the creation of a harmonious balance between physical and psychological aspects (Schuman-Olivier et al., 2020). This view is in line with the current approach to health, which focuses

ISSN: 2797-3174

more on prevention and health improvement efforts, rather than just focusing on treatment (Whitehead, 2004). However, great challenges remain in designing a truly valid, reliable, and comprehensive self health literacy assessment instrument, as the concept of health itself is complex, dynamic, and multi-dimensional.

Self health literacy is a modification of health literacy developed by Sorensen, which is the development of an instrument that focuses on self assessment of health related to the body's nutritional needs and metabolism. Health literacy can be understood as the set of cognitive and social skills that enable individuals to access, understand, process, and apply health-related information in the management of their own health (Sørensen et al., 2012). This construct is grounded in broader health literacy frameworks, notably Nutbeam's (2000) categorization into functional, interactive, and critical health literacy, and the integrated model by Sørensen et al. (2012), which emphasizes four core competencies across healthcare, disease prevention, and health promotion: access, understand, appraise, and apply. Conceptually, self-health literacy represents the operationalization of these competencies in the context of daily self-care, symptom monitoring, health-related decision-making, and the navigation of healthcare services (Sørensen et al., 2012).

Empirical studies consistently demonstrate that higher levels of health literacy, particularly when oriented toward self-care, are associated with better adherence to medication regimens, improved chronic disease self-management, and healthier lifestyle behaviors (Paasche-Orlow & Wolf, 2007; Berkman et al., 2011). For example, patients with higher health literacy tend to achieve more effective glycemic control in diabetes (Sarkar et al., 2010) and improved outcomes in chronic kidney disease self-management (Cavanaugh et al., 2008). Measurement tools such as the Health Literacy Management Scale (HeLMS) and other multidimensional instruments have been developed to capture individual competencies in accessing and applying health information in self-care contexts (Jordan et al., 2013). Nonetheless, the heterogeneity of instruments and the lack of consistent cross-cultural validation remain significant challenges, limiting comparability across populations and health systems (Jordan et al., 2013).

Strategies to enhance self-health literacy operate at both the individual and systemic levels (Brach et al., 2012). At the individual level, interventions include health education programs, digital literacy training, and self-management coaching designed to strengthen comprehension and application of health information (Sørensen et al., 2015). At the systemic level, interventions focus on simplifying health communication, redesigning services to be more user-friendly, and fostering "health-literate organizations" (Brach et al., 2012). Evidence suggests that multi component interventions combining individual skill development with structural supports yield the most sustainable outcomes (Sørensen et al., 2015). Future research priorities include establishing a clearer operational definition of self-health literacy, standardizing measurement instruments, and conducting longitudinal and intervention based studies to confirm causal pathways and assess equity impacts (Paasche-Orlow & Wolf, 2007; Berkman et al., 2011).

Several previous studies have attempted to develop self-health assessment instruments with various approaches. For example, research conducted by Post et al., (2022) developed a Rasch model-based self-health measurement tool to evaluate international spinal cord injury quality of life. The study found that the Rasch model successfully developed an instrument that has one dimension, conforms to objective measurement principles, and shows strong internal construct validity. The instrument also has a fairly good level of reliability and does not show any dependency relationship between items. However, there is a problem with the order of the available answer options, but this can be overcome by simplifying or combining some of the answer options. In addition, research by Carbó-Carreté et al. (2016) utilized a Structural Equation Modeling (SEM) approach to evaluate the link between physical activity and quality of life. The results revealed that mental health conditions and social support play an important role in shaping a person's perception of their health. However, the quantitative approach used in these studies is still not fully comprehensive in ensuring the validity and reliability of the instruments simultaneously.

In recent years, the Rasch model has been increasingly used in the development of self-assessment instruments in the health field. This model provides a strong foundation for assessing the quality of each item and adapting it to respondents' abilities. One of the main advantages of the Rasch model is its ability to ensure that the instrument developed actually measures one particular aspect consistently (unidimensional), while producing data with an interval and objective measurement scale (Bond & Fox, 2015; Sumintono & Widhiarso, 2015). Moreover, this model also helps identify items that are inappropriate or contain bias, so that the instrument becomes more accurate and fairer for various respondent backgrounds. A number of studies, including those conducted by Andrich (2018) and Linacre (2002), have shown that the Rasch model is effective in creating valid and reliable health measurement tools.

On the other hand, the Structural Equation Modeling (SEM) approach based on Partial Least Squares (PLS) has become an increasingly popular analytical method in health research, especially in testing complex relationships between latent variables, which are variables that cannot be measured directly but are represented through certain indicators. The SEM-PLS approach allows researchers to build and test models of causal relationships between various constructs while evaluating construct validity and the ability of the model

ISSN: 2797-3174

to make predictions (Hair & Alamer, 2022). In the context of self health assessment, SEM-PLS is very useful for analyzing the interaction between various factors such as lifestyle, psychological conditions, and social environment, and how these three aspects affect an individual's perception of their health. For example, the study by Seman and Tee (2018) used this approach to identify the main factors influencing health and well-being in a young adult population. The results showed that psychological and environmental factors contributed more than physical factors in shaping individuals' perceptions of their health.

Although the Rasch and SEM-PLS approaches have each been proven effective in developing self health literacy assessment instruments, the combined application of these two methods is still rarely found in research. In fact, the integration of the two can provide complementary advantages. The Rasch model is useful for ensuring the quality of the items and the suitability of the instrument with the characteristics of the respondents, while SEM-PLS helps in analyzing the relationship between variables that form the construct to be measured. Research conducted by Islam (2023) is one example that tries to combine the two approaches in designing instruments to assess health-related quality of life. The results show that this integration not only improves the accuracy and consistency of the instrument but also provides deeper insights into the various factors that influence individual health perceptions.

Based on this background, this study aims to develop and evaluating a self-health literacy assessment instrument analyze by utilizing the Rasch model and the SEM-PLS approach. Through the combination of these two methods, it is expected to produce an instrument that not only meets strict psychometric standards but is also able to describe the complexity and dynamics of self-health literacy as a whole. In addition, this study is expected to contribute both theoretically and practically in the field of public health, particularly in the development of health measurement tools that can be widely used to improve individual health awareness and quality. Thus, the findings from this study are expected to provide a strong basis for designing more effective and evidence-based health interventions.

2. Method

This research is a development of a study conducted by Kusuma et al. (2022) which focused on designing a self-health literacy assessment questionnaire and testing the internal validity and consistency of the instrument. The instrument development process was conducted in two main stages. The first stage involved identifying relevant domains or aspects of health and developing the items to be used in the instrument. This process aims to ensure that all important aspects of the concept of self-health literacy are covered thoroughly. In the second stage, the instrument was tested for content validity through expert evaluation to ensure the relevance and adequacy of the items. In addition, empirical validation was conducted by collecting data from respondents to test the extent to which the instrument can accurately measure the intended construct. Furthermore, the data results were analyzed using the Rasch model and the SEM-PLS approach to assess the quality of the items, construct validity, reliability, and the structure of the relationship between variables in the instrument. This combined approach aims to produce an instrument that is not only valid and statistically consistent but can also describe the complexity and dynamics of self-health literacy as a whole.

2.1. Item Development

Self health literacy Assessment Instrument was developed through a systematic two-stage process of content domain identification and instrument construction. In the initial stage, content domains were identified based on a comprehensive literature review, which focused on the variables to be measured. This content domain refers to topic areas relevant to the construct being measured and is obtained through an in-depth review of related literature.

The results of the literature identification refer to the health literacy dimension indicators proposed by Sørensen et al. (2012), which were then adapted into questionnaire items relevant to the health needs of students, especially related to nutrition and metabolism material. This instrument covers four main domains, namely: access to relevant health information, understanding of the information, evaluation or processing of the information, and application or use of the health information. Each item in the instrument was developed specifically based on the content domains and predefined indicators. The initial version of the instrument was then reviewed and approved by the research team to ensure content appropriateness and quality. A total of 25 items were then further tested for content validity and empirical validity in the next stage of evaluation. Details of the structure of the Health literacy Assessment instrument can be seen in Table 1 as the basis for modifying the self-health literacy instrument.

ISSN: 2797-3174

Table 1. Health Literacy Instrument Development Grid

Dimension	Aspek	Sub Aspect	Indicator	N (self health literacy instrument)
1 Dimension 1	Access/ obtain information relevant to health	Health Care	Ability to access information on medical and clinical issues	2
		Disease prevention	Ability to access information on risk factors for health	3
		Health Promotion	Ability to self-renew on determinants of health in the social and physical environment	2
Dimension 2	Understand information relevant to health	Health Care	Ability to understand medical information and derive meaning	2
		Disease prevention	Ability to understand information about risk factors and derive meaning	2
		Health Promotion	Ability to understand information about the determinants of health in the social and physical environment and derive its meaning	2
Dimension 3	Process/ appraise information relevant to health	Health Care	Ability to interpret and evaluate medical information	2
		Disease prevention	Ability to interpret and evaluate health risk factors	2
		Health Promotion	Ability to interpret and evaluate information about determinants of health in the social and physical environment	2
Dimension 4	Apply/ use information relevant to health	Health Care	Ability to make informed decisions about medical issues	2
		Disease prevention	Ability to make informed decisions about health risk factors	2
		Health Promotion	Ability to make informed decisions about determinants of health in the social and physical environment	2
	Total			25

Source by Sørensen et al (2012)

2.2. Content and Construct Validity

The content validity of an instrument can be assessed through the assessment of a group of experts consisting of experts in the field of content and educational practitioners. In this study, educational practitioners are the subject of the assessment, while content experts are professionals who have experience in related fields. The content validity assessment aims to ensure that each statement in the instrument has an appropriate and acceptable formulation, structure, consistency, and form of assessment. The experts assessed each question independently with attention to content relevance and language simplicity. Ratings were made using two four-point Likert scales, namely for relevance: 1 = not relevant, 2 = less relevant, 3 = moderately relevant, and 4 = highly relevant; and for clarity: 1 = unclear, 2 = needs revision, 3 = clear or needs little revision, and 4 = very clear. These scores were then grouped into two categories, namely relevant (scores of 3 and 4) and irrelevant (scores of 1 and 2).

The content validity of each item was calculated based on the proportion of experts who gave relevant scores (3 or 4) compared to the total number of experts. The Content Validity Index (CVI) method was used to measure the extent to which experts agreed that the instrument was content valid. CVI was calculated at the item-level (Item-CVI) and the overall instrument-level (Scale-level CVI). Item-CVI is calculated by dividing the number of experts who rated the item as highly relevant by the total number of experts. If the Item-CVI value is more than 0.79, then the item is considered feasible; a value between 0.70 and 0.79 indicates the item needs improvement; and a value below 0.70 means the item can be deleted. For the scale level, CVI can be calculated in two ways, namely based on full agreement between experts (S-CVI/UA) and based on the average of CVI values (S-CVI/Ave), where the average method is more relaxed. The difference between S-CVI/UA and S-CVI/Ave lies in the way the proportion of items that score perfectly (I-CVI = 1) is calculated. An S-CVI/UA of at least 0.8 and an S-CVI/Ave of at least 0.9 are considered to indicate excellent content validity (Devon et al., 2007), For a panel of five experts, an acceptable CVI value is 1 (Polit et al., 2007).

Based on the average Aiken V value, the tested instruments showed high validity. The value of 0.86 in the content aspect indicates that the material presented is relevant and in accordance with the purpose of the assessment, so the experts agree that the content is adequate. The construction aspect, with a score of 0.91, revealed that the structure and arrangement of the instrument were well organized, facilitating its understanding and use in the context of research or evaluation. Meanwhile, a score of 0.96 on the language aspect indicates that the use of language is very clear, precise, and free from ambiguity, thus increasing the

ISSN: 2797-3174

reliability of the instrument. Thus, since all three indicators have values above the minimum validity threshold (e.g.,≥ 0.80), the instrument can be said to be valid and suitable for use without the need for major revisions.

2.3. Participant

This survey was conducted among university pre service science from two universities in Solo-Jogja Indonesia Area: one public university (n = 173) and another private university (n = 38). Both universities were surveyed on December 13, 2024- January 3, 2025, during 100% face-to-face classes. The survey distribution process was authorized and conducted under the supervision of the relevant university institutions. Sampling utilized convenience sampling techniques to collect relevant data from the respondent group, resulting in a total of 211 (25 male, 186 female) valid responses. Respondents were included in the sample based on two main criteria, namely: (1) they were active students at the university, and (2) they were from a higher level of education in pre service science teacher. Although this study was not an experiment or quasi-experiment involving human subjects, ethical principles were applied and communicated to the respondents. These principles include: (a) freedom of participation without coercion, (b) guarantee of data confidentiality, and (c) use of the information collected for academic purposes only.

2.4. Analysis

The psychometric analysis in this study was conducted using the Rasch measurement model and analyzed through Winsteps software with a logit scale. Joint Maximum Likelihood Estimation (JMLE) was applied to run the Rasch analysis on the software. This probability-based analysis approach was used to test the fit of the responses obtained with the Rasch model pattern (Meng et al., 2023). In addition, Rasch analysis is able to identify measurement characteristics and scale structure if the model assumptions are met. This method is also useful for assessing the level of difficulty of items in a scale, as well as the ability of respondents to measure certain constructs (Bond & Fox, 2015; Meng et al., 2023; Polit et al., 2007). In the context of health and education, Rasch modeling is widely used to evaluate the psychometric quality of instruments such as scales, test items, and questionnaires (Bond & Fox, 2015; Meng et al., 2023).

To analyze the causal relationship between latent variables and evaluate the predictive ability of the model, this study used a Structural Equation Modeling (SEM) approach based on Partial Least Squares (PLS), which was operated through SEMPLS 4 software. This approach was chosen because of its ability to handle complex models with a large number of indicators and non-normal data distribution. The validity of the SEM-PLS measurement model was evaluated based on seven main criteria, namely: (1) indicator reliability, (2) construct internal consistency, (3) construct reliability, (4) convergent validity, (5) discriminant validity between constructs, (6) explained variance, and (7) standardized path coefficient. Given the limitations of Cronbach's Alpha in estimating construct reliability in reflective models, this index was not used as the main indicator in the validation process. The development of the structural model was based on the original conceptual framework developed by Hu and Schaufeli (2009), as well as referring to the correlations between dimensions in the Maslach Burnout Inventory-Student Survey (MBI-SS) instrument in previous studies. The conceptual design of this research model is presented in Figure 1.

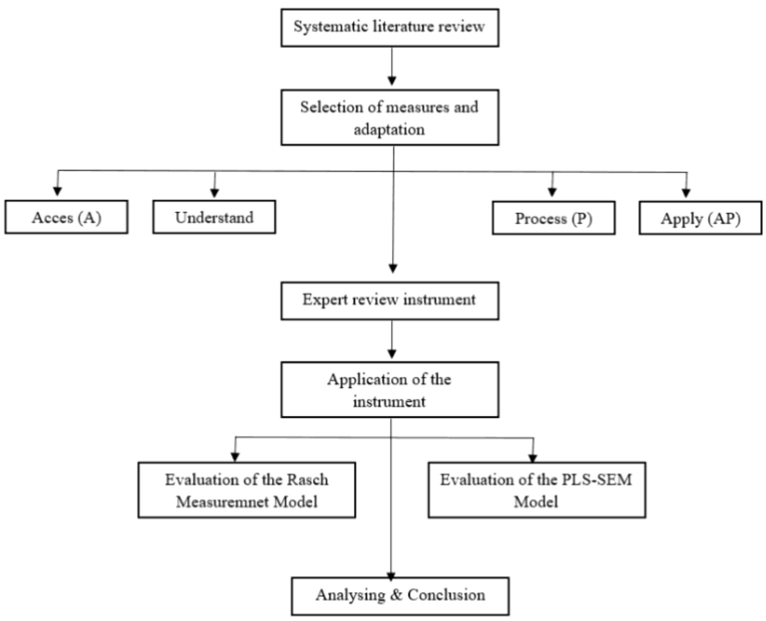

Figure 1. Research Design

ISSN: 2797-3174

3. Results and Discussion

3.1. Psychometric Analysis in Rasch Model

Evaluation of the level of fit of items to the model (fit indices) is presented in detail in Table 2.

Table 2. The Summary Statistics of Rasch Parameters for Self-Health Instrument

Criteria	Person	Item
N	211	25
Mean		
Infit MNSQ	1.02	1
Infit ZSTD	-0.19	-0.05
Outfit ZSTD	-0.2	0.03
Reliability (Rasch)	0.89	0.98
Reliability (Cronbach's alpha)	0.9	
Separation coefficient	2.8	7.35
Unidimensionality		
Raw variance by measure	41.40%	
Unexplained variance in 1st contrast	6.90%	

The data obtained shows the results of psychometric analysis on both the respondent (person) and item (item) aspects using the Rasch Model approach. The average Infit Mean Square (MNSQ) value for respondents is 1.02, which is very close to the expected value of 1. This indicates that the data has a good fit with the model. The Infit ZSTD value for respondents is -0.19, which is close to zero, indicating that the variability of responses is within reasonable limits. Similarly, the Infit ZSTD value for items of -0.05 also indicates adequate item fit to the Rasch model. The Outfit ZSTD values for both respondents and items were also close to zero, indicating that the data met the assumptions of general model fit.

The reliability level for the person model was 0.89, while for the item model it reached 0.98. This indicates high internal consistency and excellent instrument reliability. The Cronbach's Alpha value for the person model was 0.90, which reinforces the finding that the instruments have strong reliability. The separation coefficient, which describes the instrument's ability to discriminate between the respondent's ability level and the item's difficulty level, showed a value of 2.80 for respondents and 7.35 for items. This value indicates a good level of discrimination, especially in distinguishing between inter-item difficulty levels. From the aspect of unidimensionality, the raw variance explained by the main constructs is 41.40%, which indicates that most of the variance in the data can be explained by the measured constructs. Meanwhile, the unexplained variance in the 1st contrast is 6.90%, which is below the 15% threshold. This finding further strengthens that the instrument meets the assumption of unidimensionality. Overall, the analysis showed that the Self health literacy instrument has excellent psychometric qualities, including good model fit, high reliability, and a solid unidimensional structure.

3.2. Item and Person Correlation

Correlation analysis between items and respondents is therefore an important component in assessing the validity of the instrument, particularly in terms of accurately reflecting participants' ability levels and item difficulty levels. Details regarding item and respondent correlation values are presented in Table 3.

Table 3. Item Fit Statistics for Self-Health Items from the Dichotomous Rasch Analysis (n = 211).

Item	Infit		Outfit		Point Measure Correlation
	MNSQ	ZSTD	MNSQ	ZSTD	
A1	1.11	1.11	1.07	0.71	0.6
A2	0.93	-0.72	0.9	-0.83	0.48
A3	0.91	-0.96	0.91	-0.85	0.51
A4	1.06	0.7	1.08	0.78	0.47
A5	0.9	-1.12	0.96	-0.28	0.47
A6	0.87	-1.4	0.88	-1.16	0.59
A7	1.04	0.42	1.01	0.11	0.57
U1	1.22	2.2	1.3	2.84	0.53
U2	0.96	-0.36	0.94	-0.59	0.57
U3	0.81	-2.06	0.8	-2.19	0.65
U4	1	0	1.04	0.45	0.56
U5	0.89	-1.22	0.96	-0.3	0.5
U6	1.19	1.94	1.14	0.98	0.43
P1	0.94	-0.61	0.98	-0.12	0.5
P2	1.36	3.45	1.34	3.13	0.45
P3	1.06	0.62	0.93	-0.42	0.48
P4	0.79	-2.21	0.8	-2.18	0.61
P5	0.71	-3.2	0.7	-3.32	0.66

ISSN: 2797-3174

Item	Infit		Outfit		Point Measure Correlation
	MNSQ	ZSTD	MNSQ	ZSTD	
P6	0.9	-1.07	1.14	1.35	0.52
AP1	0.8	-2.19	0.79	-2.28	0.64
AP2	1.14	1.42	1.12	1.13	0.47
AP3	0.94	-0.64	0.91	-0.87	0.6
AP4	1.23	2.29	1.21	1.41	0.38
AP5	0.93	-0.71	0.97	-0.24	0.46
AP6	1.32	3.05	1.38	3.52	0.45

Based on the results of the psychometric analysis using the Rasch model, the instrument showed varying results on individual items, with most items showing a good fit with the Rasch model, while some items required improvement. In this analysis, Infit MNSQ and Outfit MNSQ were used to measure the fit of the items to the expected model, with values close to 1 indicating a good fit. Items with Infit and Outfit MNSQ close to 1, such as A1 (1.11), A4 (1.06), and U3 (0.81), indicate that they function well in measuring the intended construct and have a relatively perfect fit with the Rasch model. In addition, the high Point Measurement Correlations on items such as P5 (0.66), P4 (0.61), and U3 (0.65) indicate that the items are effective in differentiating participants based on their ability. These high correlations indicate that the items can differentiate participants well, with higher ability participants tending to select more difficult answers and vice versa. Therefore, these items can be retained in the instrument as they have been shown to have good psychometric quality. Based on these results, all statement items were declared valid and reliable. Afterward, all 25 items were found to have a good fit for the model (Figure 2).

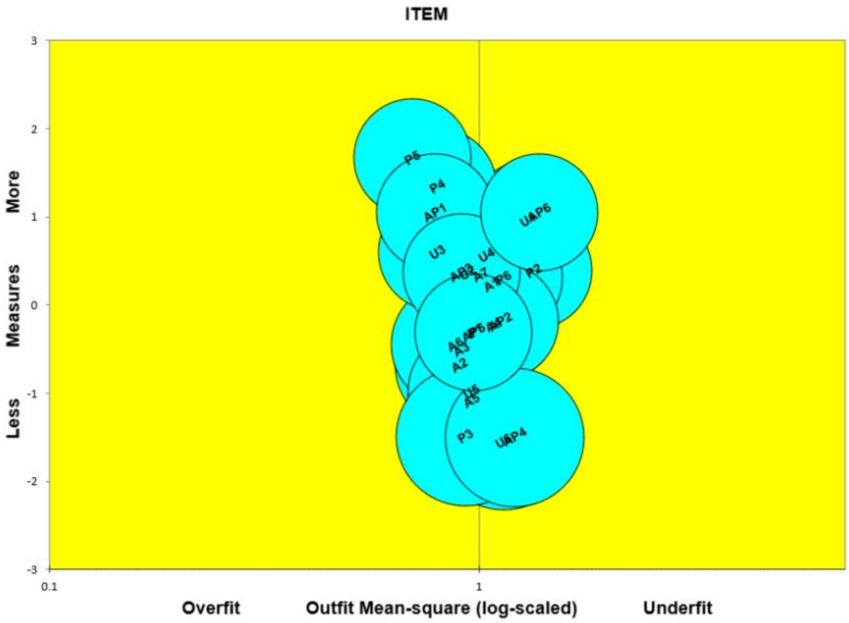

Figure 2. Bubble Diagram of Item Order Based on Infit MNSQ Value

The Y-axis shows the ability estimation based on the Joint Maximum Likelihood Estimation (JMLE) method, while the X-axis shows the mean square value of the item fit (Infit MNSQ). Items that fall into the overfit category have values greater than 1.50, while items with values between 0.50 to 1.50 are considered to be within acceptable fit limits. Each bubble in the graph represents one item, with the bubble size proportional to the standard error of the item difficulty calibration. Items that have a good fit to the model are around the center vertical line, indicating that they fit the expectations of the Rasch model.

3.3. Reliability and Internal Consistency

The reliability test results of the Self health literacy instrument showed that both person reliability and item separation were at a satisfactory level. The person separation index value of 2.8 indicates that the instrument is able to effectively distinguish individuals based on different ability levels. Meanwhile, the person reliability value of 0.89 indicates a high level of consistency in measuring individual abilities related to health aspects. This reliability figure is very good and reflects the overall stability of the instrument's measurement.

In general, a person's separation index greater than 2 indicates that the instrument is capable of distinguishing between at least three distinct levels of ability, which is a positive outcome. The person reliability index of 0.89 is also a strong result, as a reliability of 0.8 or higher typically indicates that the instrument is

ISSN: 2797-3174

consistently measuring the intended construct and provides stable results across different individuals. This suggests that the Self health literacy instrument is reliable and valid in differentiating between participants based on their health-related abilities. The item separation index value of 7.35 and item reliability of 0.98 (Hergesell, 2022) indicate that the sample size used was adequate to represent the range of item difficulty levels as shown in Figure 3. In addition, the internal consistency of the final 25 items in the Self health literacy instrument was rated as good, with a Cronbach's alpha value of 0.9. This result indicates that all items consistently measure one unidimensional construct, namely knowledge related to self-health.

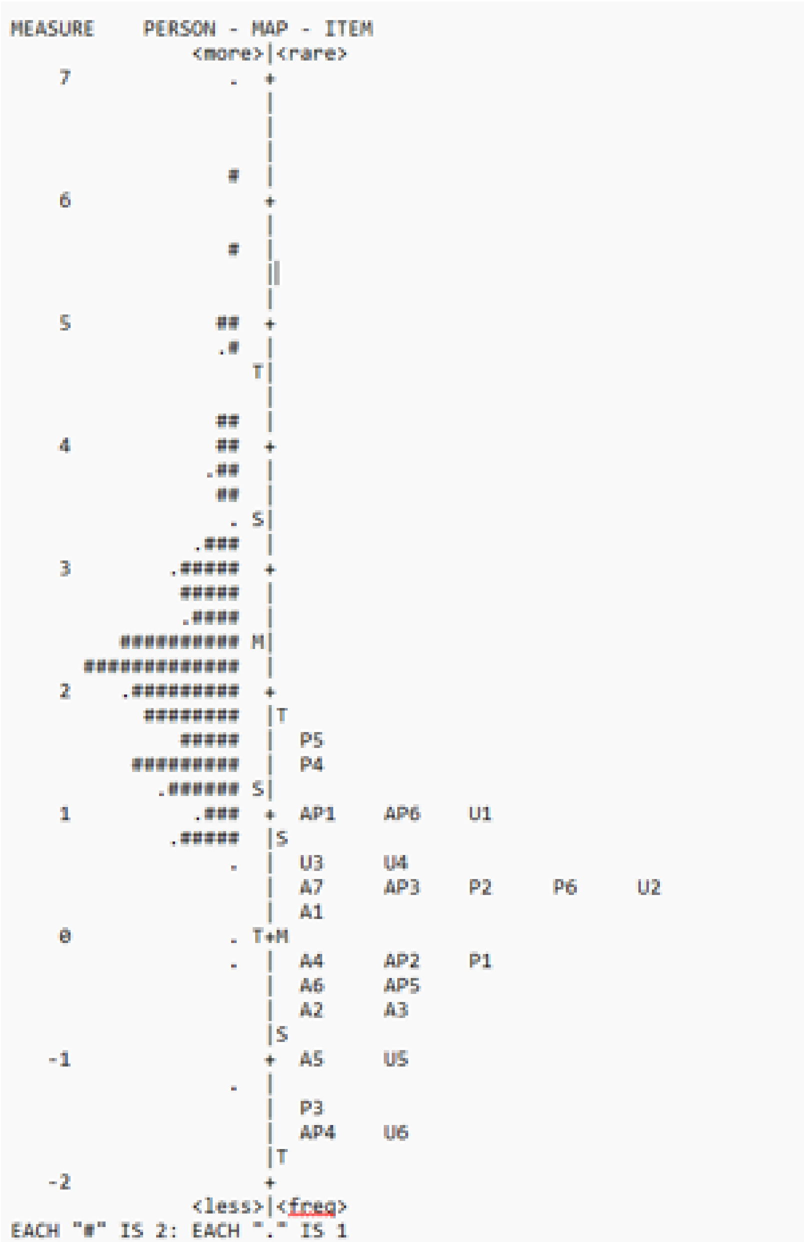

Figure 3. Wright Map Item Difficulty Level

The Figure 3 shown is the item and participant map from the Rasch analysis, which illustrates the relationship between item difficulty and participant ability level. The vertical axis of this map shows the logit, which measures both item difficulty and participant ability, with more positive logit values indicating higher item difficulty or better participant ability, while negative logit values indicate lower item difficulty or lower participant ability. The horizontal axis shows the distribution of items and participants. Each # symbol represents a participant, and the position of the participant on the logit scale indicates their ability, while the position of the item indicates its difficulty.

This map shows that most participants are around a logit value of 0, indicating that the majority have a medium level of ability. Some items, such as P5 and P4, are located at higher logit values, which means that they have a higher level of difficulty and are more suitable for participants with better ability. In contrast, items such

ISSN: 2797-3174

as A1, A2, and A3 are at lower logit levels, indicating that these items are relatively easier and suitable for participants with lower abilities. The concentrated distribution of participants in the middle of the scale also indicates the dominance of participants with moderate ability. However, there was a gap between logit values of 2 and 3, indicating that few participants had abilities in that range. This may also indicate that the instrument still lacks items of appropriate difficulty to measure ability in this range. Overall, this map provides an overview of the balance between item difficulty and participant ability. This information can be used to evaluate and improve the instrument to make it more effective in measuring the range of participants' abilities equally.

3.4. Person Ability Distribution

The analysis showed that the majority of participants had an adequate level of self health literacy. In the Rasch model framework, higher logit values reflect higher levels of knowledge. The distribution can be seen in Figure 4.

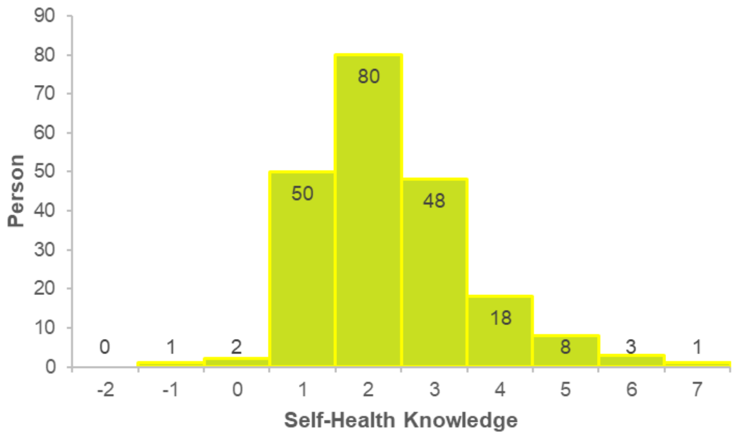

Figure 4. Summary of the Participants' Level of Knowledge About Self Health Literacy

The data showed that the majority of participants had positive logit values, with a total of 211 participants. Of the group with positive logits, there were 80 participants at level 2, 50 participants at level 1, 48 participants in the level 3 range, and 18 participants at level 4, indicating a high level of knowledge. Meanwhile, participants with lower knowledge levels consisted of 8 people at logit value 5, 3 people at logit value 6, and 1 person at logit value 7 (The distribution indicates that most participants had positive logit values, indicating a generally good level of self-health literacy. The small group at the upper end of the scale represents participants with very high levels of knowledge. This finding reflects the general pattern that the respondent population has a fairly good understanding of the concept of self-health literacy.

3.5. Internal Reliability and Validity Using SEM-PLS Model

Analysis using SmartPLS software version 4.2.7 Sarstedt et al., (2020) was conducted to obtain the value of external loadings. In reflective models, a loading value with a threshold of $\lambda \geq 0.700$ is considered ideal, but a value of $\lambda \geq 0.400$ is still acceptable as long as its presence does not reduce construct validity and the AVE (Average Variance Extracted) value remains above 0.500. If the value does not meet the criteria, the item should be removed from the model In the initial analysis, one item from the Access (A2), Understand (U6), Process (P3), and Apply (AP4) indicators was removed because its value was below the specified threshold. Furthermore, in the second analysis, items A1, A7, U5, P1, and P2 were also removed because their AVE scores still did not meet the criteria. After the deletion process, all remaining external loadings showed eligible values. The elimination process in the first and second analysis can be seen in Figure 5. Thus, the final instrument in the form of a questionnaire consisting of 16 valid items, which are divided into four main dimensions.

Some items were removed in the analysis process to ensure that the measurement tools used were valid and reliable. In the first analysis, items A2 (Access), U6 (Understand), P3 (Process), and AP4 (Apply) were removed because they did not meet the specified limits. This is in accordance with the suggestion of Hair and Alamer (2022) that items with low factor loading (usually below 0.7) should be removed for more accurate measurement. This deletion indicates that the item does not adequately describe the concept being measured or is too similar to other items in its group. In the second analysis, additional items (A1, A7, U5, P1, and P2) were also removed due to low Average Variance Extracted (AVE) values. AVE is an important measure in testing the validity of a construct, which ensures that questions within a group actually measure the same concept. This deletion is done so that the measuring instrument is of higher quality. Henseler et al. (2009) suggest this way to improve measurement accuracy.

ISSN: 2797-3174

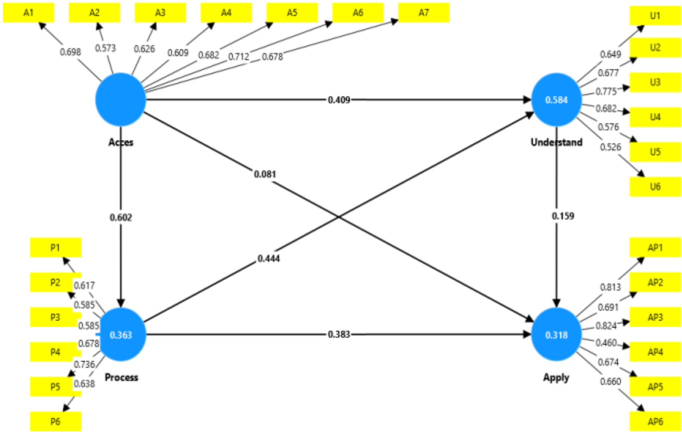

Figure 5. (a) Model Validation of Analysis 1

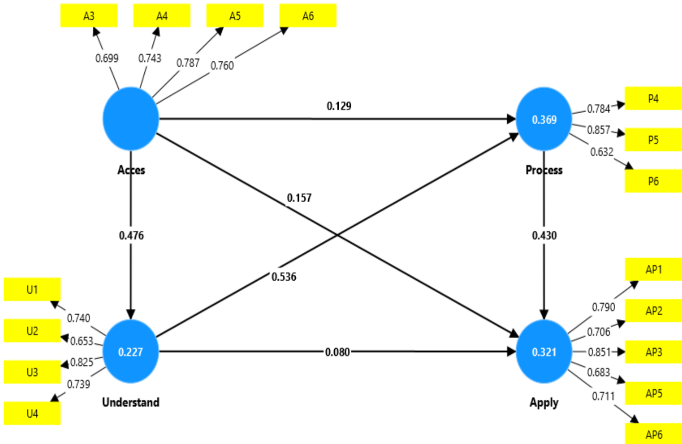

Figure 5. (b) Model Validation of Analysis 2

After the screening process, the final questionnaire consisted of 16 items divided into four categories (Access, Understand, Process, and Apply). All remaining items had good factor loadings, indicating strong validity. This is in line with the research of Chin and Newsted (1998) and Gefen et al. (2000, which emphasize the importance of high factor loading and good AVE in order for the measuring instrument to be used more reliably in future research and practice.

3.6. Internal Consistency, Reliability, and Convergent Construct Validity

The analysis results in the first and second stages showed that the four indicators met the expected value range, with rhoA ranging from 0.726 to 0.813, CR between 0.805 to 0.865, and AVE between 0.550 to 0.583 (see Table 4).

Table 4. Consistency, Reliability, and Construct Validity

Indicator	N	rhoA	Composite Reliability (CR)	AVE
Acces/obtain information relevant to health	4	0.74	0.835	0.56
Understand information relevant to health	4	0.813	0.865	0.564
Process/ appraise information relevant to health	3	0.726	0.805	0.583
Apply/use information relevant to health	5	0.734	0.829	0.55

Results from the first and second analyses showed that all four constructs Access, Understand, Process and Apply met the expected range of values, with rhoA (reliability) values ranging from 0.726 to 0.813, Composite Reliability (CR) between 0.805 and 0.865, and Average Variance Extracted (AVE) between 0.550 and 0.583 (see Table 4). These indicators are very important for evaluating the reliability and validity of the measurement model, thus providing strong evidence that the instrument developed has good consistency and robustness. The rhoA values that are above the 0.7 threshold in accordance with Dijkstra and Henseler's (2015) recommendation. indicate that the constructs have adequate internal consistency. This means that each dimension consistently measures the same concept. Furthermore, CR values that are above 0.8 reinforce the findings of the reliability of the constructs, in accordance with Hair and Alamer (2022), which states that CR values above 0.7 reflect adequate representation of latent constructs without significant measurement error.

ISSN: 2797-3174

The Average Variance Extracted (AVE) value, which is in the range of 0.550 to 0.583, is slightly above the minimum threshold of 0.5 recommended by Fornell and Larcker (1981). AVE measures the proportion of variance that can be explained by a construct compared to the variance caused by measurement error. An AVE value above 0.5 indicates that the construct is able to explain more than half of the variance of its indicators. Although this AVE value is quite close to the minimum limit, it still reflects adequate convergent validity, which means that the items in the instrument effectively represent the intended construct. This finding is in line with the results of previous research, as stated by Chin and Newsted (1998) and Gefen et al. (2000), which emphasizes the importance of achieving high reliability and validity values in structural equation models. The reported values confirm that the enhanced instrument has a good level of reliability and validity, making it suitable for further research and practical applications. However, it should be noted that while the AVE values are acceptable, improvements to these values are still possible with further refinement of the instrument items or by considering additional dimensions in future instruments.

3.7. Discriminant Validity between Constructs (HTMT)

The generally accepted HTMT threshold value is below 0.90, with a strict limit at 0.80, where values below this threshold indicate adequate discriminant validity. The results presented in Table 5 show that all HTMT values are within the acceptable range, with the highest value of 0.870, which occurs between indicators on the Apply/use information relevant to the health dimension.

Table 5. Heterotrait-monotrait Correlation Ratio (HTMT) Criterion

	1	2	3	4
1. Acces/obtain information relevant to health				
2. Understand information relevant to health	0.472			
3. Process/ appraise information relevant to health	0.583	0.737		
4. Apply/use information relevant to health	0.657	0.54	0.87	

3.8. Explained Variance (R2) and Standardized Path Coefficients (β)

In this model, the variable “Understand information relevant to health” showed an R^2 value of 0.227, which is classified as weak, while the variable “Process/appraise information relevant to health” had an R^2 value of 0.369, also in the weak category. The variable “Apply/use information relevant to health” recorded an R^2 value of 0.321, which indicates low predictive power. This finding indicates that the model still has room for improvement in explaining the variance of the construct, particularly in the context of the developed self-health literacy instrument.

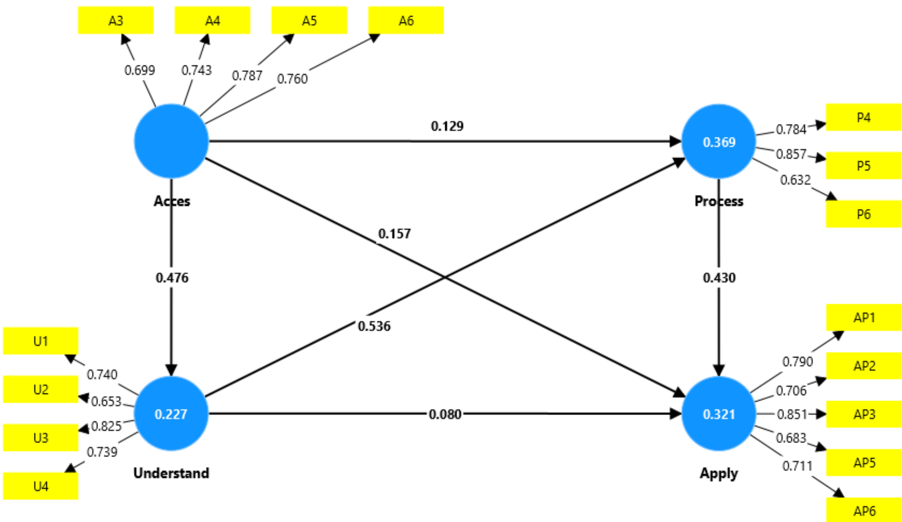

Figure 6. Nomogram of Individual Measurement Model Validation

Figure 6 also illustrates the strength of the relationship between constructs through standardized path coefficients. The expected value for this coefficient is $\beta > 0.200$, with an ideal value above 0.300 in order to be considered statistically significant. Based on the analysis, the strongest relationship was found between the Understand and Process constructs with a β value of 0.536. In contrast, the relationship between the Understand and Apply constructs has a β value of 0.080, which is low and insignificant.

4. Conclusion

The Self-Health Literacy Assessment Instrument was developed and validated using the Rasch Model and Structural Equation Modeling with Partial Least Squares (SEM-PLS). Psychometric evaluation through the Rasch

ISSN: 2797-3174

analysis demonstrated high internal consistency, with person reliability of 0.89 and item reliability of 0.98. Infit and Outfit Mean Square (MNSQ) values approximating 1.0 indicated that the majority of items fitted the model adequately, while item–measure correlations, including P5 (0.66), P4 (0.61), and U3 (0.65), suggested that the instrument was capable of effectively discriminating among participants. Further validation through SEM-PLS confirmed strong construct validity, with rhoA values exceeding 0.700 and Average Variance Extracted (AVE) values greater than 0.500 across all dimensions. 16 valid items were obtained by deleting items A1, A2, A7, U5, U6, P1, P2, P3, and AP4. These findings provided evidence of adequate convergent validity and composite reliability (CR > 0.800). Nevertheless, the R^2 values revealed varying levels of predictive power. While most dimensions showed acceptable explanatory strength, "Understand" ($R^2 = 0.227$) and "Apply" ($R^2 = 0.321$) were relatively weaker. Path coefficient analysis identified the strongest relationship between "Understand" and "Process" ($\beta = 0.536$), whereas the weakest was observed between "Understand" and "Apply" ($\beta = 0.080$). Although the instrument demonstrated satisfactory psychometric performance, refinement remains necessary to enhance predictive accuracy, particularly within the "Apply" dimension. Broader testing across diverse populations and practical implementation in health education programs are recommended to strengthen generalizability and applicability.

Author Contributions

All authors have equal contributions to the paper. All the authors have read and approved the final manuscript.

Funding

This research is funded by non-state budget funds of Universitas Sebelas Maret (non APBN UNS) under contract number: 371/UN27.22/PT.01.03/2025.

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/ or publication of this article.

Data Availability

The datasets generated during and/ or analyzed during the current study are available from the corresponding author on reasonable request.

Declaration on AI Use

The authors declare that no artificial intelligence (AI) or AI-assisted tools were used in the preparation of this manuscript.

References

Acoba, E. F. (2024). Social support and mental health: The mediating role of perceived stress. *Frontiers in Psychology, 15*, 1–12. https://doi.org/10.3389/fpsyg.2024.1330720

Andrich, D. (2018). Advances in social measurement: A Rasch measurement theory. In F. Guillemin, A. Leplège, S. Briançon, E. Spitz, & J. Coste (Eds.), *Perceived health and adaptation in chronic disease* (pp. 66–91). Routledge. https://doi.org/10.1201/9781315155074-10

Berkman, N. D., Sheridan, S. L., Donahue, K. E., Halpern, D. J., & Crotty, A. (2011). Low health literacy and health outcomes: An updated systematic review. *Annals of Internal Medicine, 155*(2), 97–107.

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). Routledge. https://doi.org/10.4324/9781315814698

Brach, C., Keller, D., Hernandez, L. M., Baur, C., Parker, R., Dreyer, B., … Schyve, P. (2012). *Ten attributes of health literate health care organizations*. National Academies Press.

Butler, L. D., Mercer, K. A., McClain-Meeder, K., Horne, D. M., & Dudley, M. (2019). Six domains of self-care: Attending to the whole person. *Journal of Human Behavior in the Social Environment, 29*(1), 107–124. https://doi.org/10.1080/10911359.2018.1482483

Carbó-Carreté, M., Guàrdia-Olmos, J., Giné, C., & Schalock, R. L. (2016). A structural equation model of the relationship between physical activity and quality of life. *International Journal of Clinical and Health Psychology, 16*(2), 147–156. https://doi.org/10.1016/j.ijchp.2015.11.001

Cavanaugh, K. L., Ellis, J., White, A. T., Avison, D., & Shulman, J. (2008). Health literacy and patient-centered medical home. *Journal of the American Board of Family Medicine, 21*(6), 576–582.

Chin, W. W., & Newsted, P. R. (1998). The partial least squares approach to structural equation modeling. In G. A. Marcoulides (Ed.), *Modern methods for business research* (pp. 295–336). Lawrence Erlbaum Associates.

ISSN: 2797-3174

Colman, D. E., Echon, R., Lemay, M. S., McDonald, J., Smith, K. R., Spencer, J., & Swift, J. K. (2016). The efficacy of self-care for graduate students in professional psychology: A meta-analysis. *Training and Education in Professional Psychology, 10*(4), 188–197. https://doi.org/10.1037/tep0000130

Corral-Verdugo, V., & Frías-Armenta, M. (2016). The sustainability of positive environments. *Environment, Development and Sustainability, 18*(4), 965–984. https://doi.org/10.1007/s10668-015-9701-7

Devon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D. J., Savoy, S. M., & Kostas-Polston, E. (2007). A psychometric toolbox for testing validity and reliability. *Journal of Nursing Scholarship, 39*(2), 155–164. https://doi.org/10.1111/j.1547-5069.2007.00161.x

Dijkstra, T. K., & Henseler, J. (2015). Consistent and asymptotically normal PLS estimators for linear structural equations. *Computational Statistics and Data Analysis, 81*, 10–23. https://doi.org/10.1016/j.csda.2014.07.008

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50.

Gefen, D., Straub, D., & Boudreau, M.-C. (2000). Structural equation modeling and regression: Guidelines for research practice. *Communications of the Association for Information Systems, 4*, Article 7. https://doi.org/10.17705/1CAIS.00407

Hair, J., & Alamer, A. (2022). Partial least squares structural equation modeling (PLS-SEM) in second language and education research. *Research Methods in Applied Linguistics, 1*(3), 1–16. https://doi.org/10.1016/j.rmal.2022.100027

Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. *Advances in International Marketing, 20*, 277–319. https://doi.org/10.1108/S1474-7979(2009)0000020014

Hergesell, A. (2022). Using Rasch analysis for scale development and refinement in tourism. *Journal of Business Research, 142*, 551–561. https://doi.org/10.1016/j.jbusres.2021.12.063

Hu, Q., & Schaufeli, W. B. (2009). The factorial validity of the Maslach Burnout Inventory–Student Survey in China. *Psychological Reports, 105*(2), 394–408. https://doi.org/10.2466/PR0.105.2.394-408

Islam, A. Y. M. A. (2023). *Applying the Rasch model and structural equation modeling to higher education: The technology satisfaction model*. Routledge. https://doi.org/10.1201/9781003384724

Jordan, J. E., Buchbinder, R., Briggs, A. M., Elsworth, G. R., Busija, L., Batterham, R., & Osborne, R. H. (2013). The Health Literacy Management Scale (HeLMS). *Patient Education and Counseling, 91*(2), 228–235.

Kissil, K., & Niño, A. (2017). Does the person-of-the-therapist training promote self-care? *Journal of Marital and Family Therapy, 43*(3), 526–536. https://doi.org/10.1111/jmft.12213

Kusuma, I. Y., Triwibowo, D. N., Pratiwi, A. D. E., & Pitaloka, D. A. E. (2022). Rasch modelling to assess psychometric validation of the KATUB-Q. *International Journal of Environmental Research and Public Health, 19*(24), Article 16753. https://doi.org/10.3390/ijerph192416753

Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement, 3*(1), 85–106.

Nutbeam, D. (2000). Health literacy as a public health goal. *Health Promotion International, 15*(3), 259–267.

Whitehead, D. (2004). Health promotion and health education. *Journal of Advanced Nursing, 47*(3), 311–320. https://doi.org/10.1111/j.1365-2648.2004.03095.x

World Health Organization. (2022). *WHO guideline on self-care interventions for health and well-being*. World Health Organization.

ISSN: 2797-3174